Article:

Tucker, R., Falvey, E., Fuller, G. orcid.org/0000-0001-8532-3500 et al. (2 more authors) (2021) Baseline SCAT performance in men and women: comparison of baseline concussion screens between 6288 elite men's and 764 women's rugby players. Clinical Journal of Sport Medicine, 31 (6). e398-e405. ISSN 1050-642X

https://doi.org/10.1097/jsm.0000000000000847

Baseline SCAT performance in men and women: comparison between 10754 elite men's and 1071 women's rugby players

Ross Tucker, Eanna Falvey, James Brown, Martin Raftery

Abstract

Background

In professional Rugby Union, mandatory annual completion of the Sports Concussion Assessment Tool (SCAT) provides reference points and clinically relevant reference limits that may be applied to enhance baseline testing and to guide return-to-play and diagnostic decisions. Women have been shown to endorse more concussion-like symptoms then men, and to outperform men in cognitive and balance tests conducted as part of concussion management assessments such as the SCAT. The differences between elite men and elite women rugby players are thus critical for effective concussion management, and this study aimed to compare SCAT performance in large cohorts of elite male (10754 players) and female (1071 players) rugby players

Results

Women endorsed significantly more symptoms, with greater symptom severity, than men (relative ratio 1.34, 95% CI 1.25 – 1.45 women vs men for any symptom). Women outperformed men in cognitive sub-modes with the exception of Immediate Memory and Delayed Recall, and made fewer balance errors than men during the mBESS. The baseline reference limits, defined as the sub-mode score that places a player in the worst-performing 5% of the cohort, were similar between men and women for all sub-modes with the exception of Concentration, Tandem Gait time and Total errors made during mBESS. Clinical reference limits, defined as sub-mode score achieved by the worst-performing 50% of the cohort, did not differ between men and women.

Conclusions

Women and men perform differently during SCAT baseline testing, though differences are small and do not affect either the baseline or clinical reference limits that identify abnormal test results for most sub-modes. The greater endorsement of symptoms by women suggests increased risk of adverse concussion outcomes, and highlights the importance of accurate evaluation of any symptom endorsement at baseline.

Keywords:

Concussion, SCAT, Rugby Union, neurological screening, concussion management

List of abbreviations

SCAT – Sports Concussion Assessment Tool

HIA - Head Injury Assessment

Declarations

Ethics approval and consent to participate

The research plan for this study was approved by the World Rugby Institutional Ethics committee (REF 19007). Players had provided written informed consent for all data gathered as part of the World Rugby Concussion management programme to be used for research in a de-identified manner

Consent for publication

Not applicable

Availability of data and material

Original participant data belongs to the players and the clubs/unions that generate such data. This may be provided upon request to third parties. World Rugby (the corresponding author) may facilitate the provision of that data, in terms of permissions and contacts, though there is not a single point of contact, since the data are generated globally from multiple teams and Unions.

Competing interests

Three of the authors (RT, EF, MH) are full-time and part-time employed by World Rugby in roles of research and medicine. GF has served as an independent advisor on a working group on concussion administered by World Rugby, for which expenses are covered. JB declares no competing interests.

Funding

There was no specific funding for the research study, though the research was supported by World Rugby, which employs some of the author group, as described in Competing Interests

Author' contributions

MR conceived the study. MR, RT and EF designed the study. RT and JB performed the analyses. All authors made substantial contributions to the study design, data processing and interpretation. RT drafted the article and all other authors revised it critically for important intellectual content. RT is the guarantor. All authors had full access to all of the data in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis.

Acknowledgements

We would like to acknowledge and sincerely thank the team doctors and medical practitioners for their help in facilitating collection of Sports Concussion Assessment Tool (SCAT) data.

Introduction

The Sports Concussion Assessment Tool (SCAT), developed after the Second International Conference on Concussion in sport {McCrory:2005im}, and subsequently revised and modified through a series of iterations to the present SCAT5 {McCrory:2017dy, Echemendia:2017kb, McCrory:2017gy}, is used in Rugby Union in various stages of its concussion management strategy {Raftery:2016ei, Fuller:2017cu}.

During match-play, an abridged version of the SCAT5 is used to guide return to play decisions after a head impact event with the potential to cause concussion. Then during subsequent diagnostic screens, complete versions of the SCAT5 support the diagnosis of concussion {Raftery:2016ei}, as part of World Rugby's head injury assessment (HIA) protocol. During these screens, symptom endorsement and the player's performance in the cognitive and balance sub-modes that make up the SCAT5 are assessed relative to a player's previously recorded baseline assessment, or, if such a baseline is absent, against clinical reference limits derived from normative baseline data obtained in a large cohort of professional rugby players {Fuller:2018eh, Fuller:2018ho}. Reference limits may also be used to indicate when a sub-mode performance is abnormal during baseline assessments, thus requiring it to be repeated to enhance its validity and resultant clinical utility.

We propose that the sub-mode reference limits used to indicate that baseline should be repeated should be set at a sub-mode score that is achieved by the worst-performing 5% of players, while a clinical reference limit, used during return-to-play and diagnostic screens, should correspond to that sub-mode score achieved by the worst-performing 50% of the cohort. This latter measure represents a more conservative concussion management strategy, reducing false negatives in clinical screening.

To date, these reference limits have been similar for female and male players. Studies have found however, that women and girls endorse more symptoms, report higher symptom severity scores {Shehata:2009db},{Covassin:2006ge, Covassin:2012co}, have different symptom profiles and improved performance in cognitive sub-modes {Shehata:2009db, Norheim:2018id, Jinguji:2012fd} and balance sub-modes {Jinguji:2012fd} compared to men and boys. These performance differences may influence the thresholds at which baseline assessments, return to play screens and diagnostic screens are deemed abnormal in women compared to men. While World Rugby recommends that all concussion testing during the HIA protocol be compared to baseline data, baseline and clinical reference limits may assist with the interpretation of diagnostic screens when baseline data are absent, and with the identification of abnormal performances during baseline screens.

Accordingly, the aim of this study was to compare SCAT5 baseline performance in large groups of professional men's and women's rugby players to identify differences in sub-mode performance, and to determine whether clinicians should apply different clinical standards to women's SCAT performance.

Methods

Study design, setting and study population

A cross sectional study was performed using data from the World Rugby Head Injury Assessment (HIA) database, which contains baseline and diagnostic concussion screen results from the professional game. In order to use the HIA process, a competition must adhere to mandatory competition player welfare standards [World Rugby Player Welfare Site] that ensures a standardised approach to concussion detection and management as well as data collection. The source population thus comprises the majority of eligible professional male players in domestic and international competitions, as well as International Women's squads that underwent mandatory baseline SCAT assessment between 2016 and 2019.

Baseline assessments

All baseline SCAT assessments were administered prior to commencement of the relevant competition season or tournament, according to methods described previously {Fuller:2018ho}. For the present analysis, we excluded baseline SCATs performed post-exercise, as well as any player who had a diagnosed concussion during the sampling period.

We chose to include players even if they had conducted multiple baseline SCATs. We recognize that this may create a learning effect due to test repetition. However, because Rugby Union requires annual baseline assessments in addition to multiple screens at the time of head impact events, most rugby players will perform multiple SCATs in their careers. Therefore, any normative reference ranges or clinical limits that are established should account for the fact that players are likely to be repeating sub-modes on multiple occasions, and so for the external validity of the data, these players with multiple tests are included, with further research studies required to quantify how sub-mode performance changes as a result of repeat testing.

Data for each sub-mode are presented as means, standard deviations, medians and the 5^{th} and 95^{th} percentile. Mean scores were compared using Mann-Whitney tests, and the null hypothesis (Men = Women) was rejected when $p < 0.004$, based on a Bonferroni correction of the original $p < 0.05$, divided by the 12 sub-domains assessed (0.05/12=0.004).

Symptoms were analysed using a Fisher's exact chi-squared analysis, comparing the proportion of each of the 22 symptoms of the SCAT5 were reported by men and women, with significance accepted when $p < 0.002$ based on a Bonferroni correction ($p < 0.05$ divided by the 22 symptoms). SCAT3 assessments were excluded from symptom analysis because we have previously shown that symptom endorsement is 32% greater using the SCAT5 than the SCAT3 (In review) likely owing to the requirement to report "trait symptoms" (how the player typically feels), compared to the "state symptoms" requested by SCAT3 {Echemendia:2017de}.

The magnitude of symptom differences between men and women was assessed by calculating a ratio (95% CI) of the proportion of women's SCAT5s in which each symptom was endorsed compared to the proportion of men's SCAT5s in which that symptom was endorsed.

Reference limits

A baseline reference limit was determined for both men and women by identifying the sub-mode score that would place the player into the worst-performing 5% of their cohort for that sub-mode. That is, the 5th or 95th percentile guided the identification of a sub-mode result that would achieve as close to 5% abnormal results as possible.

A clinical reference limit was identified similarly, but using the 50^{th} percentile to guide the identification of the sub-mode score. This clinical reference limit thus identifies the sub-mode score achieved by as close as possible to the worst-performing half of each cohort. Classifications were defined based on direction of scoring for abnormality in each sub-test, with higher symptom scores and modified Balance Error Scoring System (mBESS) errors referred to as high, and lower cognitive test performances referred to as low.

The research plan for this study was approved by the World Rugby Institutional Ethics committee (REF 19007). Players had provided written informed consent for all data gathered as part of the World Rugby Concussion management programme to be used for research in a de-identified manner.

Results

10754 SCAT assessments (4747 SCAT3 and 6008 SCAT5) were conducted in 6288 men's players, with 3660 players doing one test, 2628 performing two or more baseline SCATs during the sampling period. 1071 women's SCATs were available, comprising 263 SCAT3s and 808 SCAT5s in a total of 764 players.

Table 1 summarizes the performance in the SCAT5 sub-modes for men and women. The sample size for each sub-mode is shown, accounting for the exclusion of SCAT3 assessments for symptoms, and 5-Word lists for Immediate Memory and Delayed Recall, since these have been replaced by a 10-Word list after a ceiling effect was found to limit their utility {Echemendia:2017de, Norheim:2018id}.

	Men						Women						Mann Whitney	
	n	**Mean (SD)**	**Median**	**5th percentile**	**95th percentile**	**% perfect scores**	***n***	**Mean (SD)**	**Median**	**5th percentile**	**95th percentile**	**% perfect scores**	**Z value**	**p value**
Symptom endorsement														
Symptom number	*6008*	1.4 (2.7)	0	0	7	61%	*808*	2.2 (3.2)	1	0	9	47%	-8.913	**<0.001**
Symptom severity	*6008*	2.2 (4.7)	0	0	11	61%	*808*	3.5 (5.9)	1	0	14	47%	-8.141	**<0.001**
Cognitive sub-modes														
Orientation	*10754*	4.8 (0.4)	5	4	5	85%	*1071*	4.9 (0.4)	5	4	5	90%	-4.588	**<0.001**
Immediate Memory	*3920*	21.3 (3.7)	21	15	27	0.9%	*329*	21.8 (4.0)	22	15	29	4.0%	-2.734	0.006
Delayed Recall	*3920*	7.0 (1.9)	7	4	10	11.4%	*329*	7.2 (1.9)	7	4	10	10.9%	-1.844	0.065
Digits Backwards	*10754*	3.1 (1.0)	3	1	4	44%	*1071*	3.2 (0.9)	3	2	4	50%	-3.769	**<0.001**
Concentration	*10754*	4.0 (1.0)	4	2	5	39%	*1071*	4.1 (0.9)	4	2	5	44%	-4.456	**<0.001**
Balance sub-modes														
Tandem gait	*10195*	10.8 (2.0)	11	7.7	13.3	NA	*1035*	11.1 (1.9)	11	8	13.8	NA	-5.155	**<0.001**

Double leg balance	*10754*	0.0 (0.3)	0	0	0	97%	*1071*	0.0 (0.3)	0	0	0	99%	2.957	**0.003**
Single leg balance	*10754*	1.9 (2.0)	2	0	6	29%	*1071*	1.6 (1.8)	1	0	5	35%	5.721	**<0.001**
Tandem stance balance	*10754*	0.8 (1.3)	0	0	3	57%	*1071*	0.8 (1.5)	0	0	4	64%	4.072	**<0.001**
Total balance	*10754*	2.8 (2.8)	2	0	8	23%	*1071*	2.4 (2.7)	2	0	7	26%	5.155	**<0.001**

Table 1: Sub-mode performance in men and women

On average, women report more symptoms, with higher symptom severity than men, and outperform men in most sub-modes with the exception of Immediate Memory and Delayed Recall, where scores are similar, and Tandem Gait, which men complete faster than women (Table 1). Absolute differences in sub-mode performance are small, but statistically significant, and a greater proportion of women achieve perfect scores (no incorrect answers in cognitive sub-modes and no balance errors) more frequently than men.

On average, women were more likely to endorse symptoms (2.2 ± 2.3 symptoms in women vs 1.4 ± 2.7 for men, $p < 0.001$). Consequently, women were had a higher symptom severity score (3.5 ± 5.9 vs 2.2 ± 4.7 for women and men, respectively, $p < 0.001$). Women more frequently reported higher symptom scores, though the proportion of cases where symptoms were assessed at a score of 2 or more ("Moderate" or "Severe" on the 7-point Likert scale) was low, at 1.3% in women, compared to 0.8% for men. The 95th percentile for symptom number and severity in women was 9 and 14, respectively, compared to 7 and 11 in men (Table 1).

Table 2 shows the proportion of men and women who endorsed no symptom, any symptom and each of the 22 symptoms, while Figure 1 displays the ratio of SCAT5s in which women endorsed each symptom to SCAT5s in which men endorsed each symptom. Symptoms were grouped into physical, cognitive, vestibulo-ocular and psychological sub-groups.

60.7% of men were asymptomatic compared to 47.2% of women ($P < 0.001$, Table 2). 13 of the 22 symptoms were more likely to be endorsed by women, the most common symptoms in women being Fatigue or Low Energy (30.0% Women vs 19.4% Men), Neck Pain (20.9%

Women vs 16.2% Men), Nervous/anxious (16.7% Women vs 9.6% Men) and Trouble sleeping (15.6% Women vs 13.5% Men).

Table 2: Proportion of men's and women's baseline SCAT5s reporting each symptom

	Men	**Women**	**p-value**
	N = 6008	**N = 808**	
Asymptomatic	60.7%	47.2%	**< 0.001**
Any symptom	39.3%	52.8%	
Physical	21.2%	31.9%	**< 0.001**
Neck Pain	16.2%	20.9%	**0.001**
Headache	7.1%	12.9%	**< 0.001**
Pressure in head	6.0%	11.3%	**< 0.001**
Nausea or vomiting	1.7%	2.2%	0.336
Fatigue or low energy	19.4%	30.0%	**< 0.001**
Cognitive	18.7%	23.9%	**0.002**
Don't feel right	4.1%	4.7%	0.403
Difficulty concentrating	8.6%	13.0%	**< 0.001**
Difficultly remembering	9.4%	15.2%	**< 0.001**
Confusion	2.0%	1.5%	0.294
Drowsiness	6.1%	6.6%	0.603
Feeling slowed down	5.7%	8.4%	**0.002**
Feeling like in a fog	1.9%	2.8%	0.084
Vestibulo-Ocular	11.1%	17.9%	**< 0.001**
Dizziness	3.3%	4.1%	0.267
Blurred vision	2.8%	3.7%	0.162
Balance problems	4.2%	7.3%	**< 0.001**
Sensitivity to light	4.6%	7.8%	**< 0.001**

Sensitivity to noise	2.2%	5.0%	**< 0.001**
Psychological	20.9%	29.6%	**< 0.001**
Trouble sleeping	13.5%	15.6%	0.099
Nervous/Anxious	9.6%	16.7%	**< 0.001**
More emotional	4.5%	12.3%	**< 0.001**
Irritability	5.7%	7.5%	0.036
Sadness	2.8%	6.1%	**< 0.001**

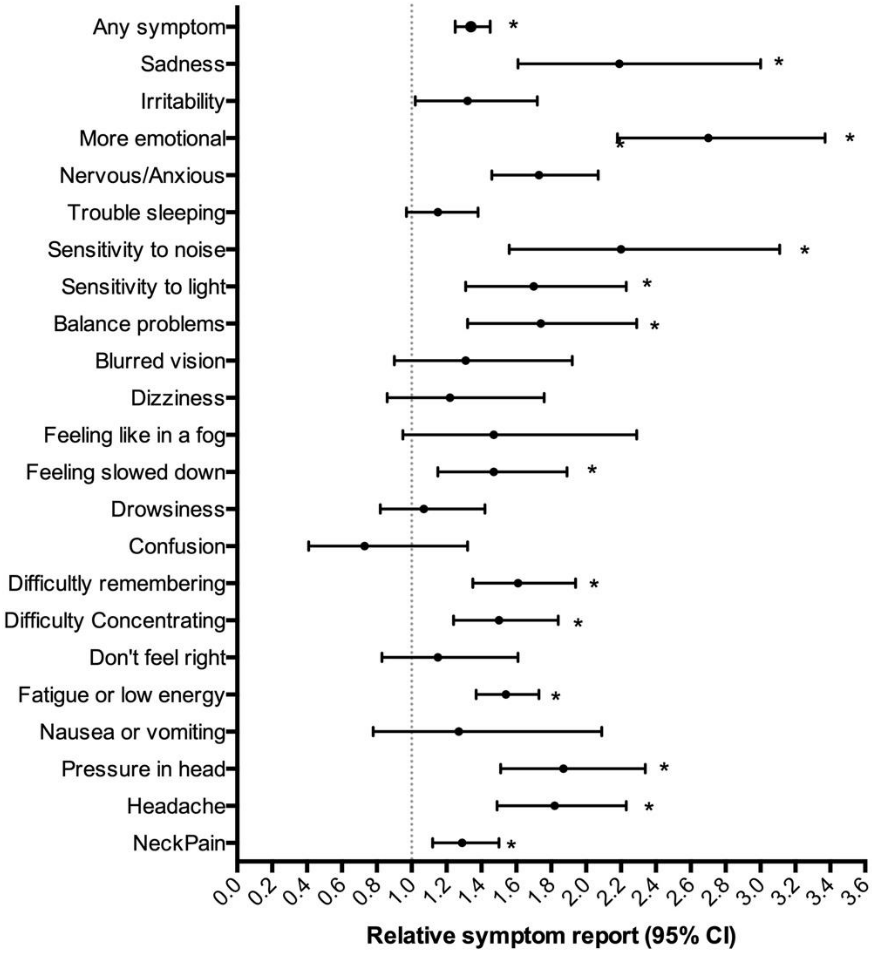

*Figure 1: Relative proportion of SCAT5s with symptom endorsement in women vs men. * significantly more likely to be endorsed in women, p < 0.002 after Bonferroni correction of the original alpha of 0.05 divided by the 22 symptoms assessed during SCAT5*

Overall, women were 34% more likely to endorse any symptom (Figure 1, M:W symptom ratio = 1.34 (1.25 – 1.45, $p < 0.001$), with relative likelihood of reporting a symptom ranging between 1.29 (Neck pain) and 2.70 (More emotional) greater for women than men in the symptoms endorsed more in women than in men (Figure 1 and Table 2).

Reference limits

Table 3 displays the derived baseline reference limits and clinical reference limits in men and women. The baseline reference limits were similar between men and women, with the exception of Concentration score (comprised of Digits Backward and Months in Reverse), Tandem Gait time and Total Balance errors. Clinical reference limits were similar with the exception of Total errors made during balance tests

Table 3: Baseline and clinical reference limits for men and women. Baseline reference limits indicate a sub-mode that requires repeat testing at baseline, and corresponds to the sub-mode score achieved by the worst-performing 5% of the population. Clinical reference limits indicate abnormal sub-mode results during clinical settings, and correspond to the worst-performing 50% of the cohort

	Baseline limit, 5%, during baseline testing		**Clinical limit, 50%, during clinical screens HIA1, HIA2, HIA3**	
	Men	**Women**	**Men**	**Women**
Cognitive sub-modes				
Orientation	3 or fewer correct answers	3 or fewer correct answers	All correct answers	All correct answers
Immediate Memory	15 or fewer correct answers	15 or fewer correct answers	21 or fewer correct answers	21 or fewer correct answers
Delayed Recall	3 or fewer correct answers	3 or fewer correct answers	7 or fewer correct answers	7 or fewer correct answers
Digits Backwards	1 or fewer correct answers	1 or fewer correct answers	3 or fewer correct answers	3 or fewer correct answers
Concentration	1 or fewer correct answers	2 or fewer correct answers	4 or fewer correct answers	4 or fewer correct answers
Balance sub-modes				
Tandem gait	13 s or slower	14 s or slower	11 s or slower	11 s or slower
Double leg balance	1 or more errors	1 or more errors	None	None
Single leg blaance	6 or more errors	6 or more errors	2 or more errors	2 or more errors
Tandem stance balance	4 or more errors	4 or more errors	None	None

Total balance	9 or more errors	8 or more errors	3 or more errors	2 or more errors

To explore baseline reference limit differences, Figure 2 shows the proportion of men and women who achieved sub-mode scores at approximately 5% for each of the three sub-modes identified as different, and for Immediate Memory for comparative purposes. For clarity, only one sub-mode score either side of the baseline reference limit is shown.

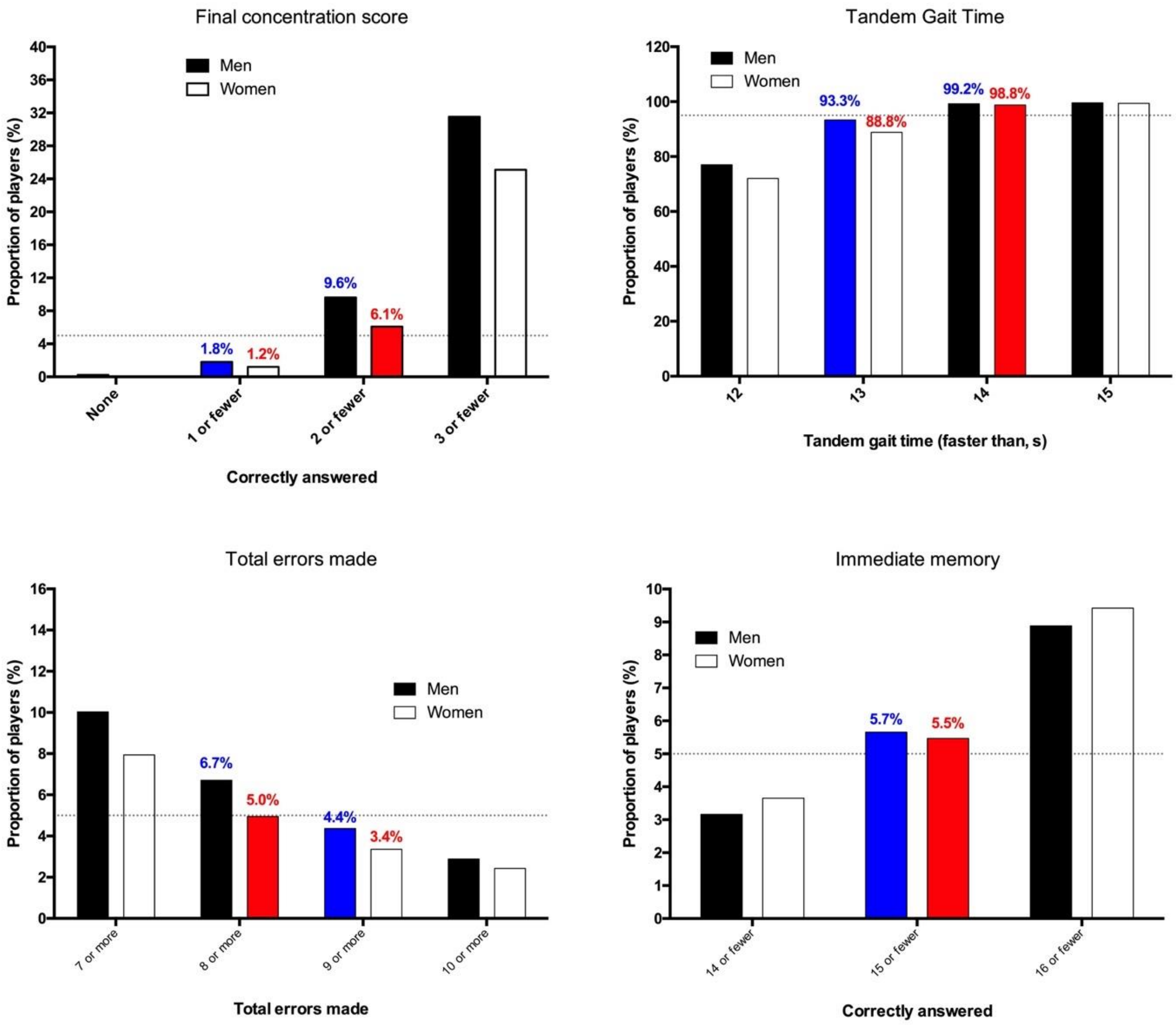

Figure 2: Proportion of men and women players achieving sub-modes scores placing them in approximately the worst-performing 5% of their respective cohorts for Final Concentration, Tandem Gait time, Total Errors made, and Immediate Memory. The baseline reference limit is identified as the sub-mode score that places a player in as close as possible to the worst-

performing 5% of the cohort. Coloured bars indicate the clinical reference limit (blue for men, red for women)

For Final concentration, the baseline reference limit was a score of 2 or fewer, achieved by 6.1% of women, compared to a baseline reference limit of 1 or fewer for men (achieved by 1.8%). 9.6% of men's players scored 2 or fewer, compared to 1.2% of women's players scoring 1 or fewer.

Tandem Gait time was significantly faster in men (Table 1), resulting in a baseline reference limit of 13s compared to 14s for women. 93.3% of men and 88.8% of women were able to complete the Tandem Gait test in under 13s, the men's baseline limit, with 98.8% of women completing the test in under 14s.

All balance mode errors were fewer in women than in men (Table 1), though a difference in baseline limit implication was found only for Total Errors, where women would be classified as abnormal at 8 or more errors, compared to 9 or more errors for men. This difference is small, however, with 6.7% of men making 8 or more total balance errors (compared to 5.0% of women), and 4.4% of men making 9 or more errors (Figure 2). Total balance errors were also different at the clinical reference limit (3 or more for men, 2 or more for women). 60.9% of men recorded 2 or more total balance errors, compared to 44.4% at 3 or more balance errors. In women, 53.3% of players made 2 or more balance errors.

Discussion

This study compared baseline SCAT performance in large cohorts of professional men's and women's rugby players. We find that women endorse more symptoms than men, report symptoms with a higher severity than men, and perform better than men in Orientation and Concentration, and balance sub-modes. Differences between women and men are however small, resulting in similar baseline reference limits for all sub-modes with the exception of Concentration, Tandem gait time and total balance errors. Clinical reference limits were similar with the exception of total balance errors.

Symptom endorsement

The greater endorsement of baseline symptoms by women, both in number and severity, is consistent with numerous previous studies {Shehata:2009db, Covassin:2007bo, Covassin:2012co, Covassin:2006ge}. One exception is Asken et al, who found no statistically significant differences in symptom severity between men and women using the SCAT3 or the SCAT5 assessment {Asken:2019ho}. However, the study included just 94 athletes compared to 6008 men and 808 women in the present study. Our work benefits from large cohorts, which results in large statistical power.

The greater endorsement of symptoms by women has implications for clinical outcomes after concussion. It has been suggested that pre-existing psychological factors may influence the incidence of all injury, particularly the severity of persistent symptoms after sports-related concussion, and perhaps the incidence of sports-related concussion itself {Trinh:2019eu}. Specifically, baseline traits of irritability, sadness, nervousness and depressive symptoms, which we found to be greater in women (Figure 1 and Table 2), predisposed athletes to worse symptomology after concussions {Trinh:2019eu, Merritt:2014db}. It has also been found that women report more symptoms and perform worse in neurocognitive tasks after concussion {Covassin:2018fg, Covassin:2012ic, Broshek:2005fg}, and suffer greater time-loss than sports-matched men after concussion {Covassin:2016kg}.

Postulated reasons for these greater adverse outcomes in women include reporting behaviours and social norms {Broshek:2005fg}, and attitude differences towards concussion that lead men to disclose concussions less often {Kerr:2016jg}. These include not wanting to be kept out of practice or matches, not wanting to let team-mates and coaches down and minimization of the seriousness of injury, possibly the result of lack of understanding {Kerr:2016jg}. The present study assesses symptoms at baseline, rather than after concussion, but the same factors may be present during annual medical assessments, in which players may downplay symptoms they fear will negatively affect their prioritized participation in the future, resulting in the lower symptom endorsement we describe among men.

With respects to management, World Rugby recommends that the team doctor review all symptoms endorsed at baseline. If these are confirmed as ‘trait’ symptoms, their cause should be investigated. Physical symptoms may have an underlying orthopaedic cause. Psychological symptoms may indicate an underlying affective disorder, and doctors are directed to the World Rugby online screening resource in the Player welfare site.

Cognitive and balance submode performances

Our second finding was that women outperform men in cognitive sub-modes with the exception of Immediate Memory and Delayed Recall, and in the mBESS sub-modes, making fewer errors. This too confirms previous research, though these previous studies have largely focused on collegiate and high school women athletes {Shehata:2009db, Norheim:2018id, Jinguji:2012fd}. The specific reasons for these differences are not clear, but may be related to years of education, innate differences between women and men, and possibly language differences between the men's and women's groups. Unfortunately, we cannot account for these differences, since the HIA database does not identify the potential characteristics that may influence cognitive and balance performance.

Reference limits

Baseline reference limits

Baseline reference limits are set at the sub-mode score achieved by the worst-performing 5% of the cohort. Effectively, this corresponds to a sub-mode score between unusually low and extremely low using the Wechsler classification {Fuller:2018ho, Hanninen:2016ew}. We find that despite the better performance of women than men in most sub-modes, there was no impact on the baseline reference limits method we have proposed to identify abnormal baseline screens, with three exceptions – Concentration, Tandem Gait and Total balance errors (Table 3). This is the result of the small size of the differences we find between men and women, which are unlikely to affect normative ranges, or cause an error in determining when a concussion has occurred, particularly given that the repeatability and inter-rater reliability of balance errors has been found to be quite low {Finnoff:2009cn}.

For final concentration, however, the difference between men and women did result in a difference in the baseline reference limit. The baseline limit for men was 1 or fewer correct answers, compared to 2 or fewer correct in women. When men and women are combined, the baseline reference limit for the entire professional rugby playing population is 2 or fewer correct answers. This is to some extent an artefact of the method used, which identifies the reference limit as the sub-mode performance that is achieved by as close to 5% of the entire cohort as possible. For Concentration, a relatively large change from 1 or fewer to 2 or fewer correct answers, as illustrated in Figure 2, results in a reference limit at a score of “1 or fewer”, achieved by 1.8% of men, rather than at 9.6% achieved by “2 or fewer”. Given the non-normal distribution of Concentration performance (Table 1), this may not warrant the application of difference baseline limits for men and women, and it may be prudent to set a limit of 2 or fewer correct answers for both men and women, even though this would result in 9.6% of men being deemed abnormal and requiring repeated baseline testing (Figure 2).

Tandem gait time was significantly faster in men, sufficient that the 5% limit for men was set at slower than 13s, and for women at slower than 14s (Table 3). The reasons for this performance difference are unknown, though the opposite finding for balance errors, where women make fewer errors than men (Table 1), suggests that a direct balance reason is not responsible. The difference may relate to foot size, where the larger average man’s foot size reduces the number of steps required to complete the test.

Clinical reference limits

Clinical reference limits are to be applied during clinical screens at HIA1, HIA2 and HIA3. We propose that the reference limit for these settings be more challenging than for repeating baseline screens, and thus identify as the sub-mode score that is achieved by the worst-performing 50% of the cohort. This more stringent clinical limit will ensure that false negatives during diagnostic screens are minimized. We have found that only total errors differs between men and women. This is true for both baseline reference limits and clinical reference limits, and is the result of the improved balance performances observed in women (Table 1).

Limitations

In each cohort, given the size and global nature of the sample, there exists a wide spread of education level, ages, languages and ethnic differences. It has previously been found that language and racial/ethnic differences do significantly impact on recall during Immediate Memory and Delayed Recall tests {Norheim:2018id} and symptom endorsement {Asken:2019hoa}, and these may have implications for concussion assessment {Norheim:2018id}. Similarly, age has been found it influence cognitive performance and symptom {Covassin:2012co, Jinguji:2012fd}, while fitness affects symptom endorsement {Naidu:2013do} and existing psychological conditions such as depression affect memory and symptoms {Covassin:2012co}. Unfortunately, we cannot yet categorize the players in our cohorts into these groups, which would allow us to explore such differences in a larger cohort than has been investigated before. It is thus a recognized limitation that our men's and women's groups may differ with respects to native English speakers, ethnic groups, age, fitness and educational background.

Conclusion

At baseline, Elite women rugby players endorse more symptoms, with greater symptom severity, than elite men players. Orientation, Concentration and balance scores are also higher in women compared to men during baseline assessment. These differences are small, and do not impact significantly on the baseline or clinical reference limits we propose to guide return to play decisions and identify abnormal sub-mode performances during baseline and diagnostic screens, with the exception of concentration, tandem gait and total balance errors. The differences between women and men for symptom report and cognitive performance, both at baseline as documented in this study and post-concussion as described in previous research, means that women may have increased risk of concussion and worse concussion outcomes. This further emphasises the importance of an accurate and valid baseline assessment, focusing in particular on symptoms, and any underlying causes for them.

References

1 McCrory P, Johnston K, Meeuwisse W, *et al.* Summary and agreement statement of the 2nd International Conference on Concussion in Sport, Prague 2004. 2005. 196–204. doi:10.1136/bjsm.2005.018614

2 McCrory P, Meeuwisse WH, Dvořák J, *et al.* 5th International Conference on Concussion in Sport (Berlin). *Br J Sports Med* 2017;**51**:837–7. doi:10.1136/bjsports-2017-097878

3 Echemendia RJ, Meeuwisse W, McCrory P, *et al.* The Sport Concussion Assessment Tool 5th Edition (SCAT5): Background and rationale. *British Journal of Sports Medicine* 2017;**51**:848–50. doi:10.1136/bjsports-2017-097506

4 McCrory P, Meeuwisse W, Dvořák J, *et al.* Consensus statement on concussion in sport-the 5th international conference on concussion in sport held in Berlin, October 2016. *British Journal of Sports Medicine* 2017;**51**:838–47. doi:10.1136/bjsports-2017-097699

5 Raftery M, Kemp S, Patricios J, *et al.* It is time to give concussion an operational definition: a 3-step process to diagnose (or rule out) concussion within 48 h of injury:

World Rugby guideline. *British Journal of Sports Medicine* 2016;**50**:642–3. doi:10.1136/bjsports-2016-095959

6 Fuller CW, Fuller GW, Kemp SPT, *et al.* Evaluation of World Rugby's concussion management process: results from Rugby World Cup 2015. *Br J Sports Med* 2017;**51**:64–9. doi:10.1136/bjsports-2016-096461

7 Fuller GW, Raftery M. Sport Concussion Assessment Tool-Fifth Edition Normative Reference Values for Professional Rugby Union Players. *Clin J Sport Med* 2018;**Publish Ahead of Print**:1. doi:10.1097/JSM.0000000000000713

8 Fuller GW, Govind O, Tucker R, *et al.* Sport concussion assessment tool-Third edition normative reference values for professional Rugby Union players. *J Sci Med Sport* 2018;**21**:347–51. doi:10.1016/j.jsams.2017.07.025

9 Shehata N, Wiley JP, Richea S, *et al.* Sport concussion assessment tool: baseline values for varsity collision sport athletes. *British Journal of Sports Medicine* 2009;**43**:730–4. doi:10.1136/bjsm.2009.059832

10 Covassin T, Swanik CB, Sachs M, *et al.* Sex differences in baseline neuropsychological function and concussion symptoms of collegiate athletes. *British Journal of Sports Medicine* 2006;**40**:923–7–discussion927. doi:10.1136/bjsm.2006.029496

11 Covassin T, Elbin RJ, Larson E, *et al.* Sex and age differences in depression and baseline sport-related concussion neurocognitive performance and symptoms. *Clin J Sport Med* 2012;**22**:98–104. doi:10.1097/JSM.0b013e31823403d2

12 Norheim N, Kissinger-Knox A, Cheatham M, *et al.* Performance of college athletes on the 10-item word list of SCAT5. *BMJ Open Sport Exerc Med* 2018;**4**:e000412. doi:10.1136/bmjsem-2018-000412

13 Jinguji TM, Bompadre V, Harmon KG, *et al.* Sport Concussion Assessment Tool-2: baseline values for high school athletes. *British Journal of Sports Medicine* 2012;**46**:365–70. doi:10.1136/bjsports-2011-090526

14 Echemendia RJ, Broglio SP, Davis GA, *et al.* What tests and measures should be added to the SCAT3 and related tests to improve their reliability, sensitivity and/or specificity in sideline concussion diagnosis? A systematic review. *Br J Sports Med* 2017;**51**:895–901. doi:10.1136/bjsports-2016-097466

15 Covassin T, Elbin RJ, Harris W, *et al.* The role of age and sex in symptoms, neurocognitive performance, and postural stability in athletes after concussion. *Am J Sports Med* 2012;**40**:1303–12. doi:10.1177/0363546512444554

16 Harrison AG, Armstrong IT, Harrison LE, *et al.* Comparing Canadian and American normative scores on the Wechsler Adult Intelligence Scale-Fourth Edition. *Archives of Clinical Neuropsychology* 2014;**29**:737–46. doi:10.1093/arclin/acu048

17 Finnoff JT, Peterson VJ, Hollman JH, *et al.* Intrarater and interrater reliability of the Balance Error Scoring System (BESS). *PM R* 2009;**1**:50–4. doi:10.1016/j.pmrj.2008.06.002

18 Covassin T, Schatz P, Swanik CB. Sex differences in neuropsychological function and post-concussion symptoms of concussed collegiate athletes. *Neurosurgery* 2007;**61**:345–50–discussion350–1. doi:10.1227/01.NEU.0000279972.95060.CB

19 Asken BM, Houck ZM, Bauer RM, *et al.* SCAT5 vs. SCAT3 Symptom Reporting Differences and Convergent Validity in Collegiate Athletes. *Archives of Clinical Neuropsychology* 2019;**32**:963–11. doi:10.1093/arclin/acz007

20 Trinh LN, Brown SM, Mulcahey MK. The Influence of Psychological Factors on the Incidence and Severity of Sports-Related Concussions: A Systematic Review. *Am J Sports Med* 2019;**42**:363546519882626. doi:10.1177/0363546519882626

21 Merritt VC, Arnett PA. Premorbid predictors of postconcussion symptoms in collegiate athletes. *Journal of Clinical and Experimental Neuropsychology* 2014;**36**:1098–111. doi:10.1080/13803395.2014.983463

22 Covassin T, Savage JL, Bretzin AC, *et al.* Sex differences in sport-related concussion long-term outcomes. *Int J Psychophysiol* 2018;**132**:9–13. doi:10.1016/j.ijpsycho.2017.09.010

23 Broshek DK, Kaushik T, Freeman JR, *et al.* Sex differences in outcome following sports-related concussion. *J Neurosurg* 2005;**102**:856–63. doi:10.3171/jns.2005.102.5.0856

24 Covassin T, Moran R, Elbin RJ. Sex Differences in Reported Concussion Injury Rates and Time Loss From Participation: An Update of the National Collegiate Athletic Association Injury Surveillance Program From 2004-2005 Through 2008-2009. *Journal of Athletic Training* 2016;**51**:189–94. doi:10.4085/1062-6050-51.3.05

25 Kerr ZY, Register-Mihalik JK, Kroshus E, *et al.* Motivations Associated With Nondisclosure of Self-Reported Concussions in Former Collegiate Athletes. *Am J Sports Med* 2016;**44**:220–5. doi:10.1177/0363546515612082

26 Asken BM, Houck ZM, Bauer RM, *et al.* SCAT5 vs. SCAT3 Symptom Reporting Differences and Convergent Validity in Collegiate Athletes. *Archives of Clinical Neuropsychology* 2019;**32**:963–11. doi:10.1093/arclin/acz007

27 Mrazik M, Naidu D, Lebrun C, *et al.* Does an individual's fitness level affect baseline concussion symptoms? *Journal of Athletic Training* 2013;**48**:654–8. doi:10.4085/1062-6050-48.3.19